

**METHOD AND ARRANGEMENT FOR SETTING UP AND UPDATING A
USER INTERFACE FOR ACCESSING INFORMATION PAGES IN A
DATA NETWORK**

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is the US National Stage of International Application No. PCT/EP2004/004148, filed April 19, 2004 and claims the benefit thereof. The International Application claims the benefits of German application No. 10320711.2 filed May 8, 2003, both applications are incorporated by reference herein in their entirety.

FIELD OF THE INVENTION

[0002] The invention relates to a method for setting up and updating a user interface for accessing information pages in a data network, and to an arrangement for carrying out the method.

SUMMARY OF THE INVENTION

[0003] Information pages in a data network, particularly the web pages on the Internet, are accessed by a user by inputting the address of the respective information page in a display program ("browser"). The address which is input is used by the display program to set up a connection to the web server on which the information page in question is stored, and as a result the information page is transmitted to the user's computer and its content is displayed there using the display program.

[0004] A displayed information page may contain "links". Links are display elements which reference (other) information pages and whose use allows these information pages to be called without the need for the addresses of these (other) information pages to be indicated to the display program, again manually.

[0005] An alternative to manual address input for accessing information pages is the use of a "search engine". A search engine can usually itself be operated using an information page and comprises at least one input field for search terms. The search engine checks a number of (ideally all) information pages to determine whether some or all search terms are contained in the content of these information pages. As search result, the search engine provides the user with an information page containing links which can be used to access some or all of the information pages found.

[0006] Another way of accessing information pages is to call them using a special user interface, the "portal page". A portal page is an information page which is provided by a portal operator and which already contains a number of links which the portal operator assumes might be useful for the user. In this case, an operator can also create various portal pages for different users or different groups of users, at least some of said portal pages being able to be changed by the user himself or by the users themselves.

[0007] Accessing information pages by virtue of the user inputting the respective address has the associated drawback that the user always needs to know the correct address of the desired information page. This problem is partly solved by virtue of the user being able to use a search engine, as described above. However, the use of search engines has the drawback that only information pages whose content includes the use of precisely those terms which the searching user has entered into the search engine's input field are found. Although when creating information pages it is possible to store additional keywords in a nondisplayed content area of the

information page, the "HTML header", which can then be found by the search engines, these keywords also need to match the search terms in order for the information page in question to be able to be found.

[0008] The drawbacks described lead the user who is looking for information to use one of the previously described portal pages, which provides a number of relevant information pages, for example for a particular subject area.

[0009] In this case, however, it has been found to be a drawback that only information pages which have already been preselected beforehand by the provider of the portal page or have been added by the user of the portal page in the course of retrospective adaptation can be accessed by means of links in this case. A portal page thus takes the user to the desired information page and hence to the desired content only if suitable links have been set up by the operator of the portal page or by a user beforehand. In this case, the number of links which can be set up is limited, merely for the sake of clarity, which means that links which are not relevant and links whose associated information page no longer exists ("dead links" or "expired links") need to be removed. In addition, on the Internet, in particular, new information pages which are potentially important to the user are frequently created afresh and need to be regularly recorded, evaluated and - depending on their importance - referenced ("linked") on the portal page by the operator of the portal page. The complexity and hence costs for maintaining a portal page thus increase proportionally to the number of links displayed.

[0010] In brief, the operation of a portal page is thus associated with a high level of predominantly manual

complexity which also needs to be put into operation on a regular basis in order to maintain the quality of the portal page.

[0011] It is the object of the invention to propose a method which reduces the complexity for setting up and updating a user interface for accessing information pages and to propose an arrangement which can be used to set up and update the user interface in a simple manner.

[0012] This object is achieved by the claims.

[0013] For the method, the solution provides for the information pages to have a respective associated content data record created for them which stores features which describe the respective content stored on the information page. To determine the selected information pages, a user data record having features which describe a sought content is created. The content data records are respectively compared with the user data record to determine the degree of match, and information pages for which the degree of match between the respective content data record and the user data record achieves and/or exceeds a previously defined threshold are displayed on the user interface by display elements as selected information pages. The effect achieved by this is that the user interface displays information pages whose content is of interest to the user.

[0014] For the arrangement, the solution provides for the information pages to have a respective content data record stored for them on the web server, said content data record describing the content which can be accessed using the respective information page, for the content data record to be able to be accessed by the portal server, for the portal server to store

a user data record which describes a content sought by the user, for the portal server to have a comparison device which compares content data records with the user data record to determine the degree of match, and for the portal server to be in a form such that information pages for which the comparison result achieves and/or exceeds a previously defined minimum match are displayed by elements on the user interface. Such an arrangement provides the user with a user interface having information pages which are matched to his needs.

[0015] Advantageous developments of the invention are described in the dependent patent claims.

[0016] If the selected information pages are checked cyclically for their accessibility and if, in cases in which one of the information pages checked can no longer be access, the display element for this information page is removed from the user interface, no inactive links are displayed on the user interface.

[0017] If the comparison is repeated at intervals of time and the result of the repeated comparison is used to update the user interface, newly added information pages with content data records are also taken into account.

[0018] If the user data record is updated, after the update the comparison is performed again, and if the new comparison result is used to determine the selected information pages again, the user interface adapts itself to suit changed requirements by the user.

[0019] Portal servers which have just been turned on can also take into account a large quantity of

information pages in the comparison if content data records are ascertained and stored by interrogating a database containing content data records and/or by means of a search, and if said content data records are used for the comparison with the user data record.

[0020] The data records are particularly simple to evaluate if each content data record and the user data record are respectively created as a structured document, and if the content data record comprises the address of its associated information page. In that case, the information page corresponding to the content data record is also particularly simple to find.

[0021] If each content data record (IDS) and the user data record are created in XML format, and if the structure of the content data records and of the user data record is respectively stipulated in a description data record, a generally known and easy-to-handle data format is used, with the use of the description data record also allowing differently structured data records to be compared with one another.

[0022] Comparisons can be made particularly easily if the same respective structure is used for the content data records and for the user data record.

[0023] The user is provided with a clear illustration if the selected information page which has been used most to date is displayed at a prioritized position on the user interface.

[0024] If the selected information page whose associated content data record has the best match with the user data record is displayed at a

prioritized position on the user interface, the user can obtain the information which is most likely to be useful to him in particularly simple fashion.

[0025] If the information pages used are web pages, and if the web pages can be accessed using addresses associated therewith, and if the web pages and the content data records are stored on at least one web server, and if the content data records are retrieved using data record addresses which are respectively associated therewith, the method may advantageously be used on the Internet or in data networks which are similar to the Internet.

[0026] The user is provided with a trusted mode of operation if the user interface used is an information page in the form of a portal page, and if the portal page is provided by a portal server.

[0027] If a proxy server is used in the data network, the proxy server reads in the information pages whose associated addresses are transmitted to it and the proxy server buffer-stores the information pages which have been read in and sends them to the user, then access to frequently required information pages is particularly effective.

[0028] Accessing information pages which have been found is particularly easy if the display elements of the selected information pages are each in the form of a link which is linked to the address of the respective selected information page, accessing the link involves the address of the associated information page being transmitted to the proxy server, and the proxy server loads this information page and sends it to the user.

[0029] If transmitting the information page to the

proxy server involves the data record address of the content data record associated with this information page likewise being transmitted to the proxy server, if the proxy server transmits this data record address to the user interface, and if the user interface loads the content data record associated with this data record address and stores it for the comparison, the volume of content data records which can be used for the comparison increases and hence so does the probability of comparison results with a high degree of match.

[0030] Exemplary embodiments of the inventive method are explained below with reference to the drawings and are used simultaneously to explain an exemplary embodiment of the inventive arrangement.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a schematic illustration of an arrangement for displaying a portal page in which a portal server is provided with information about information pages, and

Figure 2 shows a schematic illustration of the arrangement which is used to evaluate information about information pages in order to create the portal page.

DETAILED DESCRIPTION OF THE INVENTION

[0031] Figure 1 shows components of a data network (not shown in full here) which are able to interchange information with one another via the data network. In this case, there is a computer PC which is equipped with a display program for information pages, a "web browser". The web server WS is used to store information pages IS and at the same time to store

content data records IDS which are respectively associated with the information pages IS. The linking element between the web server WS and the computer PC is the proxy server PXY, which loads the information pages IS requested from the web server WS by the web browser on the computer PC from the web server, buffer-stores them and forwards them to the computer PC. Finally, the data network contains a portal server PRT which, as the user interface for the user of the computer PS, creates or holds specific information pages and sends them to the computer PC. The arrows between the individual components symbolize challenge and response messages which are interchanged between the components and which are used to explain the exemplary embodiments below. To this end, the arrows are identified by letters a) ... p).

[0032] The data network may naturally also contain a plurality of instances of all of the components; in particular, smaller data networks also contain a multiplicity of information pages IS, a plurality of computers PC and users. For reasons of clarity, however, only single instances of all components are shown in figures 1 and 2.

[0033] The web server WS, the proxy server PXY and the portal server PRT are each independent appliances in this case; it goes without saying that they may also be installed as logical entities together on a piece of hardware, however.

[0034] The data store in the web server WS stores web pages as information pages IS, said web pages being in HTML (Hypertext Markup Language) format in this case and storing various contents. A data area which is not to be displayed to the user ("HTTP header") in an information page IS stores the address of the

content data record IDS associated with this information page IS. In the present example, this content data record IDS (also called "wsmid" = "web side meta information document") is stored on the same web server WS as the associated information page IS; it goes without saying that it may also be stored on another addressable component of the data network. The content data record IDS is created by the creator of the information page IS in XML (Extensible Markup Language) format, with the structure of this content data record, that is to say details about the data fields provided in this content data record IDS, being set down in a description data record which is likewise stored on the web server WS. Such description data records are also known as DTDs (Document Type Definition) for the XML documents; they are accordingly also called "wsmid - DTD" in this case.

[0035] In the present case, the content data record has been created by the creator of the information page IS. This means that there is a high level of probability that the content data record will precisely describe the content of the information page IS. Alternatively, the content data record may also be created by a third entity, for example by a service provider specializing in this field or by an automatic device programmed specifically for this task.

[0036] The portal server PRT has a database which stores copies of content data records IDS. These copies of the content data records IDS are compared, as described later in detail, with another data record, the user data record, in order to configure the user interface for the user. The database in the portal server PRT is to this end provided with these very copies of the content data records IDS in the

course of operation. This process is described below with reference to figure 1.

[0037] The display program installed on the computer PC is used to load and display information pages IDS which are stored on the web server WS. To this end, the display program and hence the computer PC send the proxy server PXY a request message A which comprises the address of the desired information page IS. The proxy server PXY then checks whether this desired information page IS is already stored in its internal memory (cache) and, if so, whether the "expiry date" for this copy has not yet been reached. In the present example, there is not yet a copy of the information page IS in the cache memory in the proxy server PXY, which means that the latter sends a request message b), similar to the message a), to the web server WS in order to retrieve the relevant information page IS. The web server WS responds to the proxy server PXY with a response message c) which comprises the desired information page IS. As already explained, the header on the information page IS comprises the address statement for the content data record IDS associated with this information page IS. In this case, the content data record IDS, for its part, comprises the address of its associated information page IS. The proxy server PXY now uses a message d) to send the address of the content data record IDS to the portal server PRT; the content of the information page IS itself is not sent to the portal server PRT in this case, however. To send the address of the content data record IDS, an information channel reserved for this purpose is used. In data networks with a plurality of portal servers PRT, the address of the content data record can be sent to a plurality of or all portal servers PRT. The proxy server PXY now sends the information page IS loaded from the web server to the computer PC

in a response message e), and the computer PC displays this page using the display program installed thereon.

[0038] Following receipt of the message d), the portal server PRT checks whether the content data record with the data record address which has been sent is already stored in its database. This is not yet the case in the present example, which means that the portal server PRT generates a request message f) with the address of the content data record IDS and sends it to the web server WS. The latter uses a response message g) to send the content data record IDS to the portal server PRT, which evaluates the information received with this message g) in a processing process h). In this case, the received content data record IDS is conditioned such that the data fields it contains are stored as searchable objects in the database in the portal server PRT. This is done with the aid of the description of the structure of the content data record IDS, said description having been sent to the portal server PRT likewise by the web server WS in the response message g) in the form of the wsmid - DTD.

[0039] Since, as mentioned at the outset, a data network based on the invention stores and uses a multiplicity of information pages IS, the database in the portal server PRT is filled with a growing number of content data records IDS in the course of operation until - ideally - the content data records IDS of all the information pages IS are stored, provided that each information page IS also has a content data record IDS in the first place.

[0040] To provide the user of the computer PC with comprehensible access to a limited number of information pages IS, a user interface adapted to

suit the display program, a "portal page", is created for display on the display program and is displayed. This process is shown schematically in figure 2 and is explained below with reference to this figure.

[0041] The user of the computer PC activates a command in the display program in order to display a portal page as a user interface. The display program establishes from brief data interchange with the portal server PRT whether the portal server stores a portal page provided for this user. This is not the case in the present example, which means that the display program asks the user to input information about the desired content which is to be displayed for selection on the new portal page which is to be created. This can be done by the user, for example by activating selection menus or else by means of text input. From these details, the display program creates a user data record, which in this case is likewise stored in XML format and which likewise has an associated description data record (DTD). In the present example, the description data record, which is associated with the user data record, is identical to the description data record wsmid - DTD of the content data records IDS; it goes without saying that these description data records may also differ from one another, however.

[0042] The user data record now comprises, in the form of keywords, various preferences for the user, for example the language or the languages in which the information pages IS to be displayed can be written, headwords which need to be present in the title or in the content of the information pages IS, and similarly those which must not be present etc. In this case, depending on the inputs by the user, some or all fields of the user data record are filled with content. The user data record is now transmitted to

the portal server PRT from the computer PC in a request message j), and the portal server PRT compares it or the data fields contained in it with the data fields in the content data records IDS which are stored in the database in the portal server PRT. The comparison results produce a list of hits which is able to be sorted in various ways, for example according to the degree of match between all data fields, according to the statistically most frequent use of the respective information pages, age of the information page IS etc. The portal server now uses this list of hits to generate a specific information page, namely the portal page. Since the portal page is compiled afresh upon each challenge, the process is also referred to as "dynamic generation". For each successful comparison, this portal page now comprises a link to the respective information page IS found. The portal page is now transmitted from the portal server PRT to the computer PT in a response message e), and the computer PT displays it using the display program. The links displayed on the portal page are respectively provided with the title of the respective information page IS, which means that the user now selects and clicks on one or more of the displayed information pages IS from the displayed titles.

[0043] The computer PC now uses the links which have been clicked on to send a request message n), which comprises the address of the desired information page IS, to the proxy server PXY. The proxy server PXY in turn forwards the address in a further request message n) to the web server WS, which uses a response message o) to send the requested information page IS to the proxy server PXY, with the latter for its part sending the information page IS to the computer PC in the response message p). The display program on the computer PC now displays this

information page IS to the user.

[0044] If one of the information pages IS is deleted or substantially changed by the web server WS, the content of the database in the portal server PRT does not change automatically at first. This can result in the user being presented with a portal page which does not have the correct content. For this reason, the content data records IDS are deleted from the database after a preset expiry period has elapsed. Alternatively, after the expiry period has elapsed, it is also possible for a connection - not shown here - from the portal server PRT to the web server WS to be set up, within which the content data record IDS in question is transmitted to the portal server again. Changes in the content data record IDS are incorporated into the database in the portal server PRT in this case, or, if the information page IS in question and the associated content data record IDS have been deleted, the data record in question can also be deleted from the database in this case. Such checks on the database and on the content data records IDS stored therein are preferably made at low-traffic operating times in the data network.

[0045] If the user uses the display program on his computer PC to request "his" portal page, the display program checks, as described above, whether such a portal page has already been stored for the user on the portal server PRT. If this is not the case, this portal page is created afresh using the user data record which the user inputs in the course of fresh creation. If there is already a portal page, however, then the user is at this point provided with the opportunity to update the previously stored user data record, whereupon the portal page is set up afresh in line with the method described above. In this case, the user can decide whether to maintain or delete the

links which have been present to date.